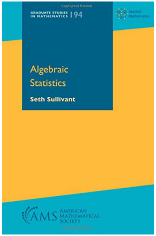
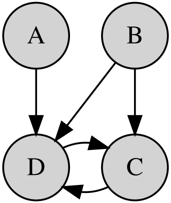
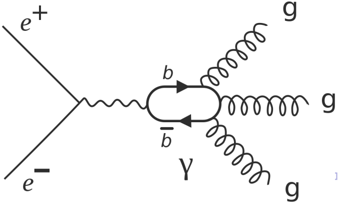
Geometric Combinatorics from Statistics and Physics

Combinatorial Coworkspace 2022

Bernd Sturmfels, MPI Leipzig







Maximum Likelihood

A *discrete statistical model* is a subset $\mathcal{M}$ of the simplex Δ_n.

The points p in Δ_n are probability distributions on the state space $\{0, 1, \ldots, n\}$. Coordinates p_i are positive and $p_0 + p_1 + \cdots + p_n = 1$.

Our data is an empirical distribution $u \in \Delta_n$. Here u_i is the fraction of samples observed to be in state i.

Maximum Likelihood

A *discrete statistical model* is a subset $\mathcal{M}$ of the simplex Δ_n.

The points p in Δ_n are probability distributions on the state space $\{0, 1, \ldots, n\}$. Coordinates p_i are positive and $p_0 + p_1 + \cdots + p_n = 1$.

Our data is an empirical distribution $u \in \Delta_n$. Here u_i is the fraction of samples observed to be in state i.

The *maximum likelihood estimator* (MLE) of $\mathcal{M}$ is the function

$$\Phi : \Delta_n \to \mathcal{M}, \;\; u \mapsto \hat{p},$$

where $\hat{p} \in \mathcal{M}$ is the maximizer of the *log-likelihood function*

$$p \;\mapsto\; \sum_{i=0}^{n} u_i \cdot \log(p_i).$$

Key Point: *If $\mathcal{M}$ is a variety then Φ is an algebraic function. The algebraic degree of Φ is the maximum likelihood degree of $\mathcal{M}$.*

A picture from [ASCB '05]

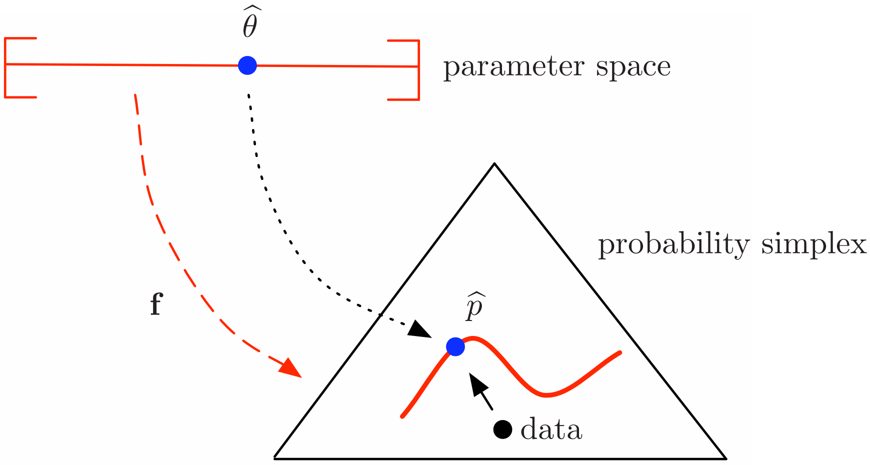
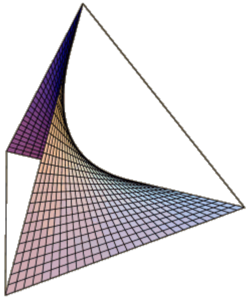


Fig. 3.2. The geometry of maximum likelihood estimation.

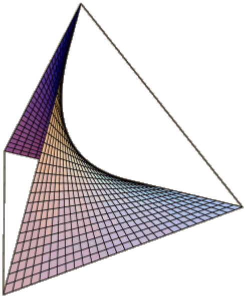
Independence



Let $n = 3$ and consider two binary random variables. The independence model $\mathcal{M}$ is a *surface in the tetrahedron* Δ_3.

Points in the model are positive rank one 2×2 matrices $\begin{bmatrix} p_0 & p_1 \\ p_2 & p_3 \end{bmatrix}$ whose entries sum to one. **Data** is a 2×2 integer matrix u.

Independence



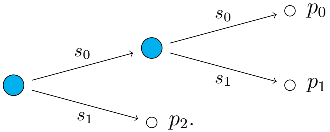
Let $n = 3$ and consider two binary random variables. The independence model $\mathcal{M}$ is a *surface in the tetrahedron* Δ_3.

Points in the model are positive rank one 2×2 matrices $\begin{bmatrix} p_0 & p_1 \\ p_2 & p_3 \end{bmatrix}$ whose entries sum to one. **Data** is a 2×2 integer matrix u. The sample size is $|u| = u_0 + u_1 + u_2 + u_3$. The empirical distribution is $\frac{1}{|u|} u$.

The ML degree is 1 because the MLE is a rational function:

$$\hat{p}_0 = |u|^{-2}(u_0+u_1)(u_0+u_2), \qquad \hat{p}_1 = |u|^{-2}(u_0+u_1)(u_1+u_3),$$
$$\hat{p}_2 = |u|^{-2}(u_2+u_3)(u_0+u_2), \qquad \hat{p}_3 = |u|^{-2}(u_2+u_3)(u_1+u_3).$$

Probability Tree



Experiment: *Flip a biased coin. If it shows heads, flip it again*.
The model $\mathcal{M}$ also has ML degree 1. It is the parametric curve

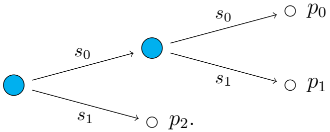
$$\Delta_1 \to \Delta_2\,,\ (s_0, s_1) \mapsto (s_0^2, s_0 s_1, s_1) \qquad \text{where } s_0, s_1 > 0 \text{ and } s_0 + s_1 = 1.$$

Implicit representation:

$$\mathcal{M} \ = \ \{\, (p_0, p_1, p_2) \in \Delta_2 \, : \, p_0 p_2 = (p_0 + p_1) p_1 \,\}.$$

Perform experiments and record outcomes in the data vector (u_0, u_1, u_2) .

Probability Tree



Experiment: *Flip a biased coin. If it shows heads, flip it again*.
The model $\mathcal{M}$ also has ML degree 1. It is the parametric curve

$$\Delta_1 \to \Delta_2\,,\ (s_0, s_1) \mapsto (s_0^2, s_0 s_1, s_1) \qquad \text{where } s_0, s_1 > 0 \text{ and } s_0 + s_1 = 1.$$

Implicit representation:

$$\mathcal{M} \ = \ \{\, (p_0, p_1, p_2) \in \Delta_2 \ : \ p_0 p_2 = (p_0 + p_1) p_1 \,\}.$$

Perform experiments and record outcomes in the data vector (u_0, u_1, u_2) .
Estimated parameters are empirical frequency of heads and tails:

$$\hat{s}_0 \ = \ \frac{2u_0 + u_1}{2u_0 + 2u_1 + u_2} \quad \text{and} \quad \hat{s}_1 \ = \ \frac{u_1 + u_2}{2u_0 + 2u_1 + u_2}.$$

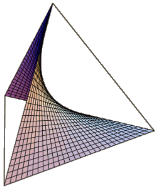
MLE is given by alternating products of positive linear forms:

$$(\hat{p}_0, \hat{p}_1, \hat{p}_2) \ = \ \left(\frac{(2u_0 + u_1)^2}{(2u_0+2u_1+u_2)^2}, \frac{(2u_0+u_1)(u_1+u_2)}{(2u_0 + 2u_1 + u_2)^2}, \frac{u_1 + u_2}{2u_0+2u_1+u_2} \right)$$

How to be Rational?

Analogy: Let $\mathcal{M} \subset \mathbb{R}^n$ and $\Phi : \mathbb{R}^n \to \mathcal{M}$ its Euclidean nearest point map. Then Φ is a rational function if and only if $\mathcal{M}$ is a linear space.

Question: For statistical models $\mathcal{M} \subset \Delta_n$, role of Φ is played by MLE. *Which $\mathcal{M}$ play the role of linear spaces? Which $\mathcal{M}$ have ML degree 1?*



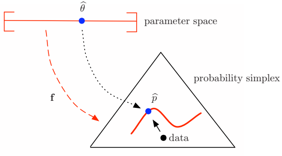
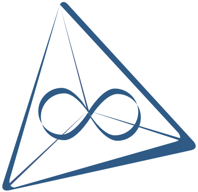


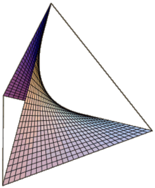
Fig. 3.2. The geometry of maximum likelihood estimation.



How to be Rational?

Analogy: Let $\mathcal{M} \subset \mathbb{R}^n$ and $\Phi : \mathbb{R}^n \to \mathcal{M}$ its Euclidean nearest point map. Then Φ is a rational function if and only if $\mathcal{M}$ is a linear space.

Question: For statistical models $\mathcal{M} \subset \Delta_n$, role of Φ is played by MLE. *Which $\mathcal{M}$ play the role of linear spaces? Which $\mathcal{M}$ have ML degree 1?*



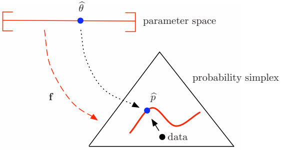
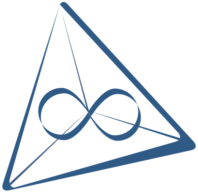


Fig. 3.2. The geometry of maximum likelihood estimation.



Solution given by [June Huh: *Varieties with maximum likelihood degree one*, J. Algebraic Statistics, 2014]

refined by [Eliana Duarte, Orlando Marigliano, B.St.: *Discrete statistical models with rational maximum likelihood estimator*, Bernoulli, 2021]

First Exercise

Do Now

The following **model** has four outcomes: 0, 1, 2, or 3 heads.

Flip a biased coin. If it shows tails, stop.
Otherwise flip it again. If it shows tails, stop.
Otherwise flip it again. Stop. Record the number of heads.

This **experiment** is carried out N times. The **data** are summarized in a vector $(u_0, u_1, u_2, u_3) \in \mathbb{N}^4$ with $u_0 + u_1 + u_2 + u_3 = N$.

1. The model is a curve in the tetrahedron Δ_3, and hence in the projective space $\mathbb{P}^3$. Determine its homogeneous prime ideal.
2. Write the MLE $(\hat{p}_0, \hat{p}_1, \hat{p}_2, \hat{p}_3)$ explicitly in terms of the data.

First Exercise

Do Now

The following **model** has four outcomes: 0, 1, 2, or 3 heads.

Flip a biased coin. If it shows tails, stop.
Otherwise flip it again. If it shows tails, stop.
Otherwise flip it again. Stop. Record the number of heads.

This **experiment** is carried out N times. The **data** are summarized in a vector $(u_0, u_1, u_2, u_3) \in \mathbb{N}^4$ with $u_0 + u_1 + u_2 + u_3 = N$.

1. The model is a curve in the tetrahedron Δ_3, and hence in the projective space $\mathbb{P}^3$. Determine its homogeneous prime ideal.
2. Write the MLE $(\hat{p}_0, \hat{p}_1, \hat{p}_2, \hat{p}_3)$ explicitly in terms of the data.

$$2 \times 2\text{-minors of} \quad \begin{pmatrix} p_0 & p_1 + p_2 & p_2 \\ p_1 & p_3 - p_1 & p_3 - p_2 \end{pmatrix}$$

$$\hat{p}_0 = \frac{(3u_0+2u_1+u_2)^3}{(3u_0+3u_1+2u_2+u_3)^3} \qquad \hat{p}_1 = \frac{(3u_0+2u_1+u_2)^2(u_1+u_2+u_3)}{(3u_0+3u_1+2u_2+u_3)^3}$$

$$\hat{p}_2 = \frac{(3u_0+2u_1+u_2)(u_1+u_2+u_3)}{(3u_0+3u_1+2u_2+u_3)^2} \qquad \hat{p}_3 = \frac{u_1+u_2+u_3}{3u_0+3u_1+2u_2+u_3}.$$

Projective Varieties

Let $\mathcal{M}$ be a variety in $\mathbb{P}^n$ with coordinates $(p_0 : p_1 : \cdots : p_n)$.
For $u \in \mathbb{Z}^{n+1}$, the likelihood function is the rational function

$$L_u : \mathcal{M} \dashrightarrow \mathbb{C}, \quad p \mapsto \frac{p_0^{u_0} p_1^{u_1} \cdots p_n^{u_n}}{(p_0 + p_1 + \cdots + p_n)^{|u|}}$$

Let $\mathcal{H}$ be the arrangement of $n+2$ hyperplanes in this formula.

Theorem

The number of complex critical points of L_u is independent of u, provided u is generic. It is the ML degree of the model $\mathcal{M} \cap \Delta_n$, and it equals the signed Euler characteristic of $\mathcal{M}\backslash\mathcal{H}$, provided this very affine variety is smooth. Otherwise, $\mathrm{MLdegree}(\mathcal{M})$ equals the Chern-Schwartz-MacPherson class of $\mathcal{M}\backslash\mathcal{H}$, provided ....

[F. Catanese, S. Hoșten, A. Khetan, B. St: *The maximum likelihood degree*, American J. Math, 2006]

[J. Huh: *The maximum likelihood degree of a very affine variety*, Compositio Math, 2013]

[J. Huh, B. St.: *Likelihood geometry*, Combinatorial algebraic geometry, Springer LNM, 2014]

[J. Rodriguez, B. Wang: *Computing Euler obstruction functions using maximum likelihood degrees*, IMRN, 2020]

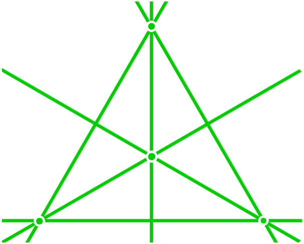
Hyperplane Arrangements

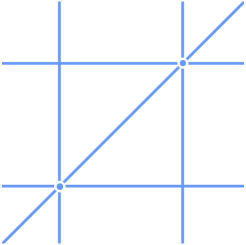
Theorem (Varchenko 1995)

If $\mathcal{M}$ is a linear space then $\mathrm{MLdegree}(\mathcal{M})$ is the number of bounded regions in the hyperplane arrangement $\mathcal{M} \cap \mathcal{H}$ in $\mathbb{R}^n$, where $\{p_0 + p_1 + \cdots + p_n = 0\}$ is the hyperplane at infinity. For positive u, every complex solution is real, one per bounded region.

Combinatorics and Linear Algebra: matroid, characteristic polynomial, Zaslavsky,....

[T. Brysiewicz, H. Eble, L. Kühne: *Enumerating chambers of hyperplane arrangements with symmetry, May 2021*]





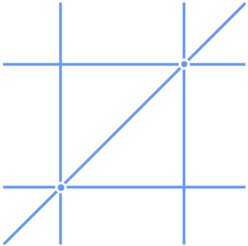
Second Exercise

Do Now

Fix positive integers $a, b, c, d, e, f, g, h, i$, and consider the following polynomial function in three complex variables:

$$x_1^a x_2^b x_3^c (x_1 - x_2)^d (x_1 - x_3)^e (x_2 - x_3)^f (1 - x_1)^g (1 - x_2)^h (1 - x_3)^i$$

How many critical points does this function have?



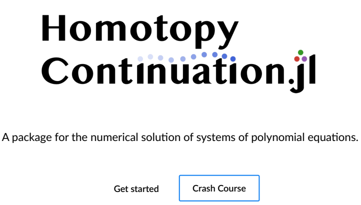
From Linear to Nonlinear

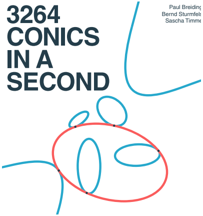
Theorem (Varchenko 1995)

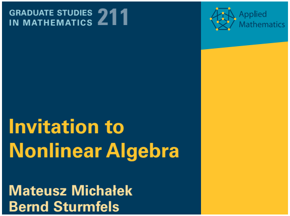
If $\mathcal{M}$ is a linear space then $\mathrm{MLdegree}(\mathcal{M})$ is the number of bounded regions in the hyperplane arrangement $\mathcal{M} \cap \mathcal{H}$ in $\mathbb{R}^n$, where $\{p_0 + p_1 + \cdots + p_n = 0\}$ is the hyperplane at infinity. For positive u, every complex solution is real, one per bounded region.

Q: What if $\mathcal{M}$ is not linear? How to compute critical points of L_u?

A: Use **Numerical Methods** from **Nonlinear Algebra**.







Six Particles on a Line

The moduli space $M_{0,6} = \mathrm{Gr}(2,6)^\circ/(\mathbb{C}^*)^6$ is a very affine variety of dimension 3. It embeds into $\mathbb{P}^8\backslash\mathcal{H}$ by the 2×2 minors of

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & x_1 & x_2 & x_3 & 1 \end{bmatrix}.$$

The scaled minors sum to 1:

$$\begin{array}{ccc} p_{23} = 5x_1/9, & p_{24} = x_2/3, & p_{25} = x_3/9, \\ p_{34} = (x_2 - x_1)/9, & p_{35} = (x_3 - x_1)/9, & p_{45} = (x_3 - x_2)/9, \\ p_{36} = (1 - x_1)/3, & p_{46} = (1 - x_2)/3, & p_{56} = (1 - x_3)/3. \end{array}$$

Given u_{ij}, we seek the critical points of the log-likelihood function

$$\log(L_u) = \sum u_{ij} \log p_{ij}(x)$$

This is the potential in the CHY model. Its derivatives are the scattering equations. Mandelstam invariants u_{ij} represent data.

[F. Cachazo, S. He, E. Yuan: *Scattering equations and Kawai-Lewellen-Tye orthogonality*, Physical Review D, 2014]

[F. Cachazo, N. Early, A. Guevara, S. Mizera: *Scattering equations: from projective spaces to tropical Grassmannians*, J. High Energy Physics, 2019]

Joy of Numerics

For certain positive integers u_{ij}, we find the six critical points:

$$
\begin{aligned}
\hat{x}_1 &= 0.2400432759291, & \hat{x}_2 &= 0.5081722067398, & \hat{x}_3 &= 0.7770058668172;\\
x_1 &= 0.2234375508553, & x_2 &= 0.8435430486816, & x_3 &= 0.5187063898083;\\
x_1 &= 0.4819677264510, & x_2 &= 0.2355452408806, & x_3 &= 0.7811156798859;\\
x_1 &= 0.6182779262092, & x_2 &= 0.8519744569452, & x_3 &= 0.1559925583741;\\
x_1 &= 0.8619960607096, & x_2 &= 0.2176050433439, & x_3 &= 0.4532389470048;\\
x_1 &= 0.8631924172503, & x_2 &= 0.5786694562520, & x_3 &= 0.1579601163959.
\end{aligned}
$$

The first solution gives the MLE

$$
\begin{aligned}
&\hat{p}_{23} = 0.13336, \quad \hat{p}_{24} = 0.16939, \quad \hat{p}_{25} = 0.08633, \quad \hat{p}_{34} = 0.02979, \quad \hat{p}_{35} = 0.05966,\\
&\hat{p}_{36} = 0.25332, \quad \hat{p}_{45} = 0.02987, \quad \hat{p}_{46} = 0.16394, \quad \hat{p}_{56} = 0.07433.
\end{aligned}
$$

We note:

- all six solutions are real,
- for each permutation ijk, one solution has $0 < x_i < x_j < x_k < 1$,
- six bounded regions in the arrangement of planes $\{p_{ij}(x) = 0\}$.

[B St and Simon Telen: *Likelihood equations and scattering amplitudes*, Algebraic Statistics, 2021]

CHY Scattering Equations

The moduli space has dimension $m-3$ and is *very affine*:

$$M_{0,m} \;=\; \mathrm{Gr}(2,m)^{\circ}/(\mathbb{C}^*)^m.$$

$$\begin{bmatrix} 0 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & x_1 & x_2 & \cdots & x_{m-3} & 1 \end{bmatrix}$$

This is a **linear** statistical model in $\mathbb{P}^n$, with $n+1 = m(m-3)/2$ states.

Mandelstam invariants serve as the data u. CHY scattering equations characterize critical points of the log-likelihood function. By Varchenko, the ML degree is $(m-3)!$. For positive u, all critical points are real.

CHY Scattering Equations

The moduli space has dimension $m-3$ and is *very affine*:

$$M_{0,m} \;=\; \mathrm{Gr}(2,m)^{\circ}/(\mathbb{C}^*)^m.$$

$$\begin{bmatrix} 0 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & x_1 & x_2 & \cdots & x_{m-3} & 1 \end{bmatrix}$$

This is a **linear** statistical model in $\mathbb{P}^n$, with $n+1 = m(m-3)/2$ states.

Mandelstam invariants serve as the data u. CHY scattering equations characterize critical points of the log-likelihood function. By Varchenko, the ML degree is $(m-3)!$. For positive u, all critical points are real.

Certified numerical solutions found with `HomotopyContinuation.jl`:

m	$n+1$	$(m-3)!$	$t_{\mathbb{C}}$	$t_{\mathbb{R}}$	t_{cert}
10	35	5040	0.75	0.28	0.5
11	44	40320	13.4	3.4	4.0
12	54	362880	124.6	43.7	45.0
13	65	3628800	2141.5	578.2	1178.0

[Paul Breiding, Kemal Rose, Sascha Timme: *Certifying zeros of polynomial systems using interval arithmetic*, 2020]

Higher Dimensions

The moduli space for m points in general position in $\mathbb{P}^{k-1}$,

$$X(k,m) \;:=\; \mathrm{Gr}(k,m)^{\circ}/(\mathbb{C}^*)^m,$$

is smooth and very affine of dimension $(k-1)(m-k-1)$.

Euler characteristic is the ML degree of the *CEGM model*.

$\underline{k=3}$: m points in $\mathbb{P}^2$, no three collinear. Numerical computation yields:

m	$n+1$	ML degree	$t_{\mathbb{C}}$	t_{cert}
6	14	26	0.02	0.01
7	28	1272	0.35	0.19
8	48	188112	70.03	47.71
9	75	74570400	*last*	*slide*

These ML degrees were known, thanks to Cachazo et al. and Thomas Lam (matroids, finite fields, Weil). *We confirmed them.*

Questions: Can all critical points be real? What about larger k, m?

Third Exercise

Do Now

The 4-dimensional very affine variety $X(3,6)$ is given by matrices

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 & x_1 & x_3 \\ 1 & 0 & 0 & 1 & x_2 & x_4 \end{bmatrix},$$

where $x_1, x_2, x_3, x_4 \in \mathbb{C}$ and all 3×3-minors all nonzero.

1. What are the fibers of the map $X(3,6) \to X(3,5)$ given by deleting the last column? Find their Euler characteristic.
2. Show that $X(3,5) \simeq X(2,5)$. Find its Euler characteristic.
3. Determine the Euler characteristic of $X(3,6)$.

Third Exercise

Do Now

The 4-dimensional very affine variety $X(3,6)$ is given by matrices

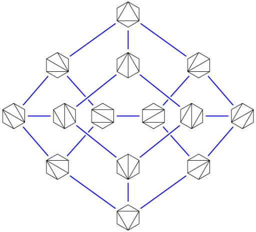
$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 & x_1 & x_3 \\ 1 & 0 & 0 & 1 & x_2 & x_4 \end{bmatrix},$$

where $x_1, x_2, x_3, x_4 \in \mathbb{C}$ and all 3×3-minors all nonzero.

1. What are the fibers of the map $X(3,6) \to X(3,5)$ given by deleting the last column? Find their Euler characteristic.
2. Show that $X(3,5) \simeq X(2,5)$. Find its Euler characteristic.
3. Determine the Euler characteristic of $X(3,6)$.

$$26 = 13 \times 2$$

Amplitudes and Positive Geometries



CEGM amplitudes are rational functions in the data u:

$$\frac{1}{u_{12}u_{34}u_{56}} + \frac{1}{u_{12}u_{56}u_{123}} + \frac{1}{u_{23}u_{56}u_{123}} + \frac{1}{u_{23}u_{56}u_{234}} + \frac{1}{u_{34}u_{56}u_{234}} + \frac{1}{u_{16}u_{23}u_{45}} + \frac{1}{u_{12}u_{34}u_{345}}$$
$$+ \frac{1}{u_{12}u_{45}u_{123}} + \frac{1}{u_{12}u_{45}u_{345}} + \frac{1}{u_{16}u_{23}u_{234}} + \frac{1}{u_{16}u_{34}u_{234}} + \frac{1}{u_{16}u_{34}u_{345}} + \frac{1}{u_{16}u_{45}u_{345}} + \frac{1}{u_{23}u_{45}u_{123}}$$

They can be evaluated either combinatorially, or numerically, by summing residues over all critical points of the likelihood function.

Inspired by

[N. Arkani-Hamed, Y. Bai and T. Lam: *Positive geometries and canonical forms*, J High Energy Physics, 2017]

[N. Arkani-Hamed, S. He and T. Lam: *Stringy canonical forms*, J High Energy Physics, 2021]

... we defined positive statistical models and their amplitudes.

These include all *toric models, linear models, ML degree one models*, ...

Frequentist to Bayesian

Statisticians study marginal likelihood integrals

$$\int_\Theta p_0(x)^{u_0} p_1(x)^{u_1} \cdots p_n(x)^{u_n} \mu(x) \mathrm{d}x$$

for parametrized models $\Theta \to \mathcal{M} \subset \mathbb{P}^n$.

The amplitude is a limit of a certain transformation.

Theorem

The amplitude of a positive model $\mathcal{M}$ is a rational function of the data $u_0, \ldots, u_n$. It equals the volume of the dual polytope, and can be computed from the toric Hessian of the log-likelihood:

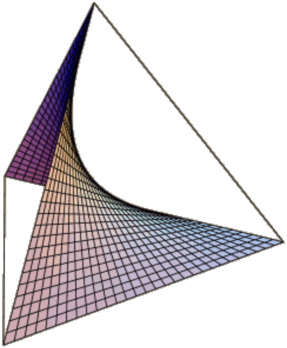
$$\mathrm{amplitude}(\mathcal{M}) \quad = \quad \sum_{\xi \in \mathrm{Crit}(L_u)} \det(H_{L_u}(\xi))^{-1}.$$

Aside: In algebraic statistics, this relates to moment varieties:

[K. Kohn, B. Shapiro, B. St: *Moment varieties of measures on polytopes*, Annali della Scuola Normale, 2020]

[K. Kohn, K. Ranestad: *Projective geometry of Wachspress coordinates*, Foundations of Computational Math,2020]

Toric Model



Back to two binary random variables ($n = 3$).
The independence model is the surface $\mathcal{M} = \mathbb{P}^1 \times \mathbb{P}^1$ in $\mathbb{P}^3$.

The relevant polytope is the square $[0,1]^2$ times $|u|$. Translate by $u = (u_2 + u_3, u_1 + u_3)$, dualize, and measure the area, to get

$$\mathrm{amplitude}(\mathcal{M}) = \frac{(u_0 + u_1 + u_2 + u_3)^2}{(u_0+u_1)(u_2+u_3)(u_0+u_2)(u_1+u_3)}.$$

A Paper with Six Authors

[Daniele Agostini, Taylor Brysiewicz, Claudia Fevola, Lukas Kühne, BSt and Simon Telen: **Likelihood Degenerations**]

We work over the *Puiseux series field* $\mathbb{R}\{\{t\}\}$ and use its valuation.

In English: We introduce a small positive parameter t with $t \to 0$.

The data u have their coordinates in $\mathbb{R}\{\{t\}\}$:

$$u_i(t) \quad = \quad \alpha_i t^{w_i} + \text{higher order terms}$$

Critical points $\hat{p}$ have coordinates in the algebraic closure $\mathbb{C}\{\{t\}\}$:

$$\hat{p}_j(t) \quad = \quad \beta_j t^{q_j} + \text{higher order terms}$$

Tropical MLE problem: Given the tropical data $w = (w_0, \ldots, w_n)$ compute the tropical critical points $q = (q_0, \ldots, q_n)$ in terms of w.

[D. Maclagan and B. St: *Introduction to Tropical Geometry*, AMS, 2015]

Numerics meets Tropics: Happiness both ways

Exercise: Solve the quintic $x^5 - 5x^4 - 4x^3 + 20x^2 - tx + 7t^4 = 0$.

Solution: The five zeros are $x(t) =$

$$2 - \frac{1}{24}t - \frac{5}{6912}t^2 + \cdots, \quad -2 - \frac{1}{56}t + \frac{25}{87808}t^2 + \cdots, \quad 5 + \frac{1}{105}t - \frac{71}{1157625}t^2,$$

$$\frac{1}{20}t + \frac{1}{2000}t^2 - \frac{5599967}{800000}t^3 + \cdots, \qquad 7t^3 + 980t^5 + 274400t^7 + \cdots.$$

$$x(t) \sim t^0, t^0, t^0, t^1, t^3, \qquad \mathrm{val}(x(t)) \ = \ 0, 0, 0, 1, 3$$

Numerics meets Tropics: Happiness both ways

Exercise: Solve the quintic $x^5 - 5x^4 - 4x^3 + 20x^2 - tx + 7t^4 = 0$.

Solution: The five zeros are $x(t) =$

$$2 - \frac{1}{24}t - \frac{5}{6912}t^2 + \cdots, \quad -2 - \frac{1}{56}t + \frac{25}{87808}t^2 + \cdots, \quad 5 + \frac{1}{105}t - \frac{71}{1157625}t^2,$$

$$\frac{1}{20}t + \frac{1}{2000}t^2 - \frac{5599967}{800000}t^3 + \cdots, \qquad 7t^3 + 980t^5 + 274400t^7 + \cdots.$$

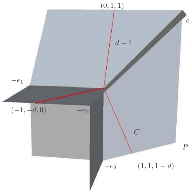
$$x(t) \sim t^0, t^0, t^0, t^1, t^3, \qquad \mathrm{val}(x(t)) = 0, 0, 0, 1, 3$$

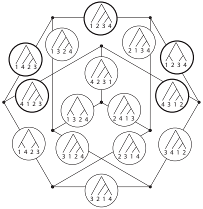
From **T** to **N**: Identify tropical solutions (cf. Newton polygon). Use this to build a homotopy for numerical solving over $\mathbb{R}$ or $\mathbb{C}$.

From **N** to **T**: Use advanced numerical tools (e.g. monodromy) to solve equations for $t = 10^{-3}, 10^{-4}, 10^{-5}, \ldots$. From the resulting numerical data over $\mathbb{C}$, learn the tropical solutions.

We applied **both directions** to MLE and scattering equations.

Linear Models





matroids, Bergman fans, nbc bases, ...

Let X be a linear subspace of $\mathbb{P}^n$, viewed as a statistical model. We tropicalize both X and its orthogonal complement $X^\perp$.

Theorem

If the tropical data vector w is sufficiently generic then the following intersection consists of $\mathrm{MLdegree}(X)$ many distinct points. These are the tropical critical points

$$\hat{q} \ \in \ \mathrm{trop}(X) \ \cap \ (w - \mathrm{trop}(X^\perp)).$$

Corollary

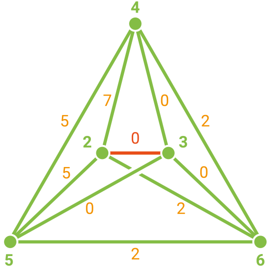
If X is general then $\hat{q} = w_0 e_{n+1} + \sum_{i \in I} w_i e_i$, where I runs over all d-sets in $\{1, \ldots, n\}$. These are the $\binom{n}{d}$ tropical critical points.

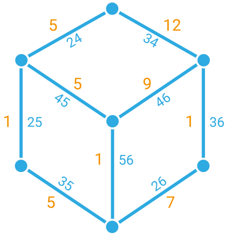
Tropical CHY Scattering

CHY model $M_{0,6}$ is a 3-dim'l linear space X in $\mathbb{P}^8$, giving an arrangement of 9 planes in $\mathbb{R}^3$, with six bounded regions. Fix

$$w \;=\; (w_{24}, w_{25}, \ldots, w_{45}) \;=\; (12, 6, 9, 12, 5, 1, 10, 11, 3)$$

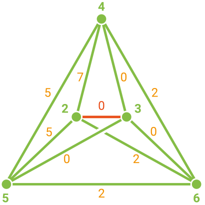
One of the six tropical critical points is $\hat{q} = (7, 5, 2, 0, 0, 0, 5, 2, 2)$:

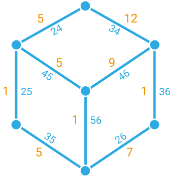




Solutions are decompositions $w = \hat{q} + (w - \hat{q})$ whose summands respect circuits and cocircuits. Each gives a small arc in one region that converges with the given rates to a vertex of the arrangement.

Fourth Exercise





Consider the six critical points $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ of

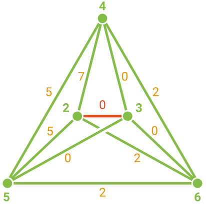
$$x_1^a x_2^b x_3^c (x_1 - x_2)^d (x_1 - x_3)^e (x_2 - x_3)^f (1 - x_1)^g (1 - x_2)^h (1 - x_3)^i$$

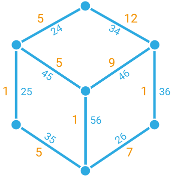
where $\big(\mathrm{val}(a), \mathrm{val}(b), \ldots, \mathrm{val}(i)\big) \;=\; (12, 6, 9, 12, 5, 1, 10, 11, 3).$

Compute all six tropical critical points: We already know

$$\hat{q} \;=\; \big(\mathrm{val}(\hat{x}_1), \mathrm{val}(\hat{x}_2), \ldots, \mathrm{val}(1 - \hat{x}_3)\big) \;=\; (7, 5, 2, 0, 0, 0, 5, 2, 2).$$

Fourth Exercise





Consider the six critical points $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ of

$$x_1^a x_2^b x_3^c (x_1 - x_2)^d (x_1 - x_3)^e (x_2 - x_3)^f (1 - x_1)^g (1 - x_2)^h (1 - x_3)^i$$

where $\big(\mathrm{val}(a), \mathrm{val}(b), \ldots, \mathrm{val}(i)\big) = (12, 6, 9, 12, 5, 1, 10, 11, 3)$.

Compute all six tropical critical points: We already know

$$\hat{q} = \big(\mathrm{val}(\hat{x}_1), \mathrm{val}(\hat{x}_2), \ldots, \mathrm{val}(1 - \hat{x}_3)\big) = (7, 5, 2, 0, 0, 0, 5, 2, 2).$$

The six tropical critical points $\hat{q}$ are

$$(0,0,8,4,2,0,2,0,0), \quad (0,5,2,2,0,0,0,0,2), \quad (1,0,8,0,2,0,0,1,0),$$
$$(2,5,2,0,0,0,2,3,2), \quad (7,5,2,0,0,0,5,2,2), \quad (9,0,8,0,2,0,0,8,0).$$

Soft Limits

Cachazo, Umbert and Zhang (2020) studied tropical MLE for the model $X(k,m)$ with very specific *tropical Mandelstam invariants*

$$w_I \;=\; \begin{cases} 1 & \text{if } m \in I \\ 0 & \text{if } m \notin I \end{cases}$$

indexed by k-sets $I \subset \{1,\ldots,m\}$.

Solutions are *regular* if they tropicalize to $\hat{q} = 0$; otherwise *singular*.

Regular solutions arise from generic fibers $F_{k,m}$ of the map

$$X(k,m+1) \;\longrightarrow\; X(k,m).$$

The $F_{k,m}$ are generic discriminantal hyperplane arrangements.

We count their bounded regions. Singular solutions are more subtle. They come from special geometric loci in $X(k,m)$.

We compute these loci using numerical–tropical happiness.

Discriminantal Arrangements

Theorem

For fixed dimension k, the number of bounded regions of $F_{k,m}$ is a polynomial in m of degree $(k-1)^2$. For $k = 3,4,5,6,7$, it equals

$$\frac{m-3}{8}(m^3 - 3m^2 + 2m - 8)$$

$$\frac{m-4}{1296}(m^8 - 5m^7 - 68m^6 + 772m^5 - 3299m^4 + 7153m^3 - 7650m^2 + 3096m + 1296)$$

$$\frac{m-5}{7962624}(m^{15} - 19m^{14} + 165m^{13} - 1687m^{12} + 14947m^{11} + 20847m^{10} - 1883209m^9 + 19445731m^8 - 105532464m^7 + 347718184m^6 - 704585488m^5 + 815190576m^4 - 398830464m^3 - 84195072m^2 + 112637952m - 7962624)$$

$$\frac{m-6}{2985984000000}(m^{24} - 44m^{23} + 911m^{22} - 11784m^{21} + 97541m^{20} - 336204m^{19} - 4467549m^{18} + 115776456m^{17} - 1593224629m^{16} + 13128969276m^{15} - 19383488419m^{14} - 1059764682264m^{13} + 16113981947031m^{12} - 136378934149764m^{11} + 803680447423841m^{10} - 3541838991169704m^9 + 12070676668677656m^8 - 32308966820835264m^7 + 67944291044051216m^6 - 110339489042552704m^5 + 133034610370502400m^4 - 111021306363648000m^3 + 56477160852480000m^2 - 12997485711360000m + 2985984000000)$$

$$\frac{m-7}{1003061300428800000000}(m^{35} - 83m^{34} + 3304m^{33} - 83972m^{32} + 1530340m^{31} + \cdots - 1003061300428800000000)$$

[H. Koizumi, Y. Numata, A. Takemura: *On intersection lattices of hyperplane arrangements generated by generic points*, Annals of Combinatorics, 2012]

From Soft Limits to Hard Facts

Theorem

The ML degree of $X(4,8)$ equals **5,211,816**.

Proof involves solving likelihood equations for all subproblems representing singular solutions. Massive numerical computation with lots of combinatorial bookkeeping. Blueprint for future tropical MLE.

We also verified that the ML degree of $X(3,9)$ equals

$$\mathbf{74{,}570{,}400} \;=\; 188112 \cdot 205 \,+\, 420 \cdot 81040 \,+\, 105 \cdot 18768.$$

Our main theoretical result is a proof of the CUY conjecture on *special loci in* $X(3,m)$. [§5: Fibrations of points and lines in the plane]

From Soft Limits to Hard Facts

Theorem

The ML degree of $X(4,8)$ equals **5,211,816**.

Proof involves solving likelihood equations for all subproblems representing singular solutions. Massive numerical computation with lots of combinatorial bookkeeping. Blueprint for future tropical MLE.

We also verified that the ML degree of $X(3,9)$ equals

$$\mathbf{74{,}570{,}400} \;=\; 188112 \cdot 205 \,+\, 420 \cdot 81040 \,+\, 105 \cdot 18768.$$

Our main theoretical result is a proof of the CUY conjecture on *special loci in $X(3,m)$*. [§5: Fibrations of points and lines in the plane]

Conclusion: *Particle Physics guides us towards new computational paradigms for likelihood inference in Algebraic Statistics. On route, we are having fun with Combinatorics and Algebraic Geometry.*

Thanks for Listening